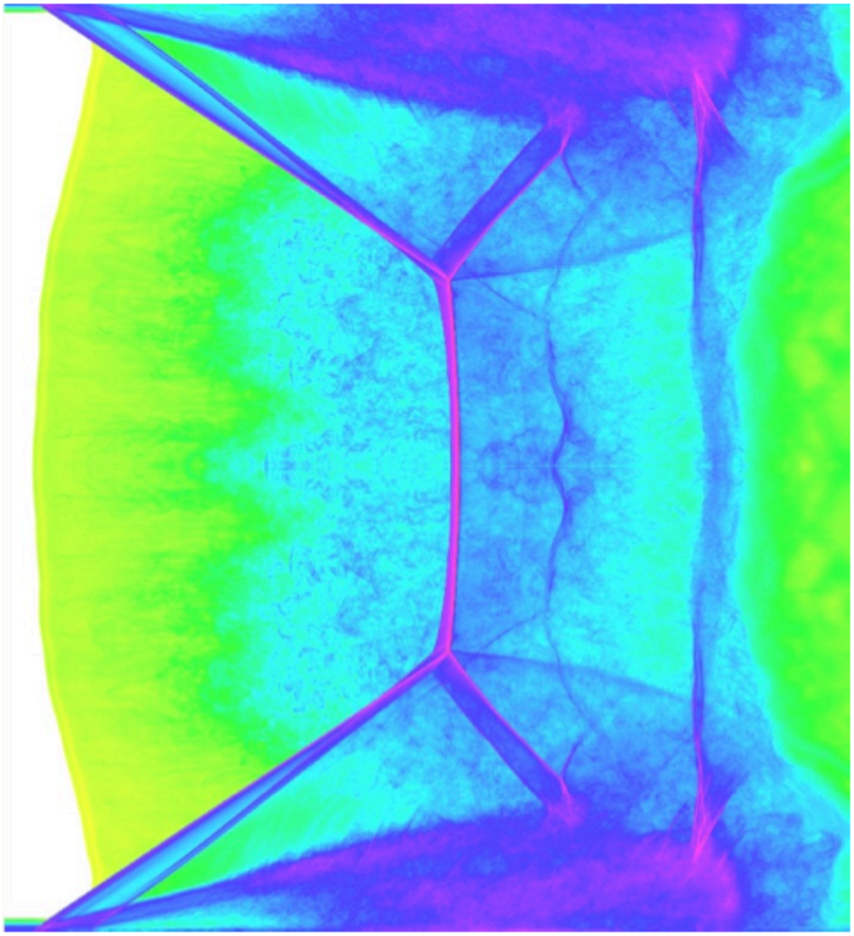
Tuning High Speed Combustion and Detonation code for Mira

Alexei Khokhlov, University of Chicago
Joanna Austin, University of Illinois
Andrew Knisely, University of Illinois
Charles Bacon, Argonne National Laboratory
Ben Clifford, Argonne National Laboratory
Marta García, Argonne National Laboratory

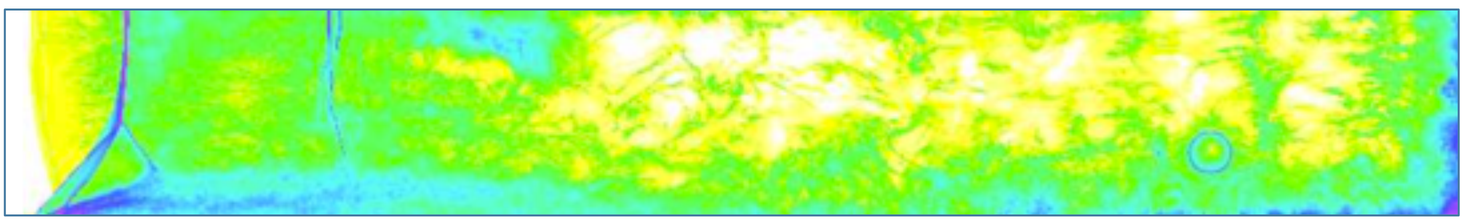
Overview

- The Science
- Setup and code features
- Scaling challenges
- Simulations
- Current timings
- Remaining challenges



3D Navier-Stokes first-principles DNS of a Mach=3 reflected shock bifurcation in a $2H_2$-O_2 mixture in a square channel. Pseudo-Schlieren image.

The Science …



The Science

PI Name:	Alexei Khokhlov
Institution:	University of Chicago
Co-PI:	Joanna Austin (U of I), Charles Bacon (ANL)
Allocation Program:	ESP
Allocation Hours:	150 Million
Year:	2010 – 2013
Research domain:	Chemistry

Direct Numerical Simulations of the deflagration-to-detonation transition (DDT) in hydrogen-oxygen gaseous mixtures for hydrogen safety (funded by DOE ASCR).

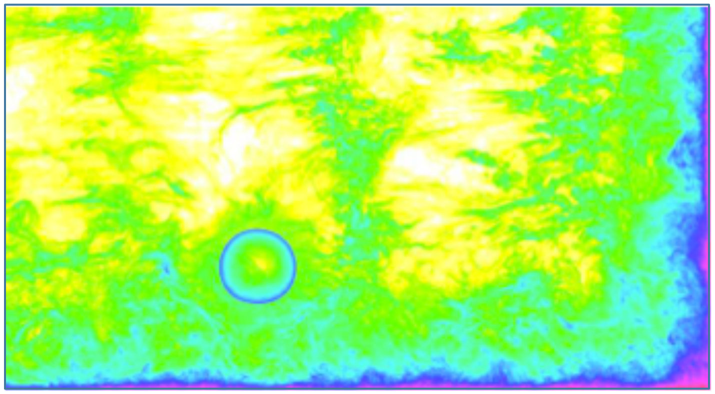
Steps:

I. Simulations of reflected shock tube experiments in CO_2 (~10M)

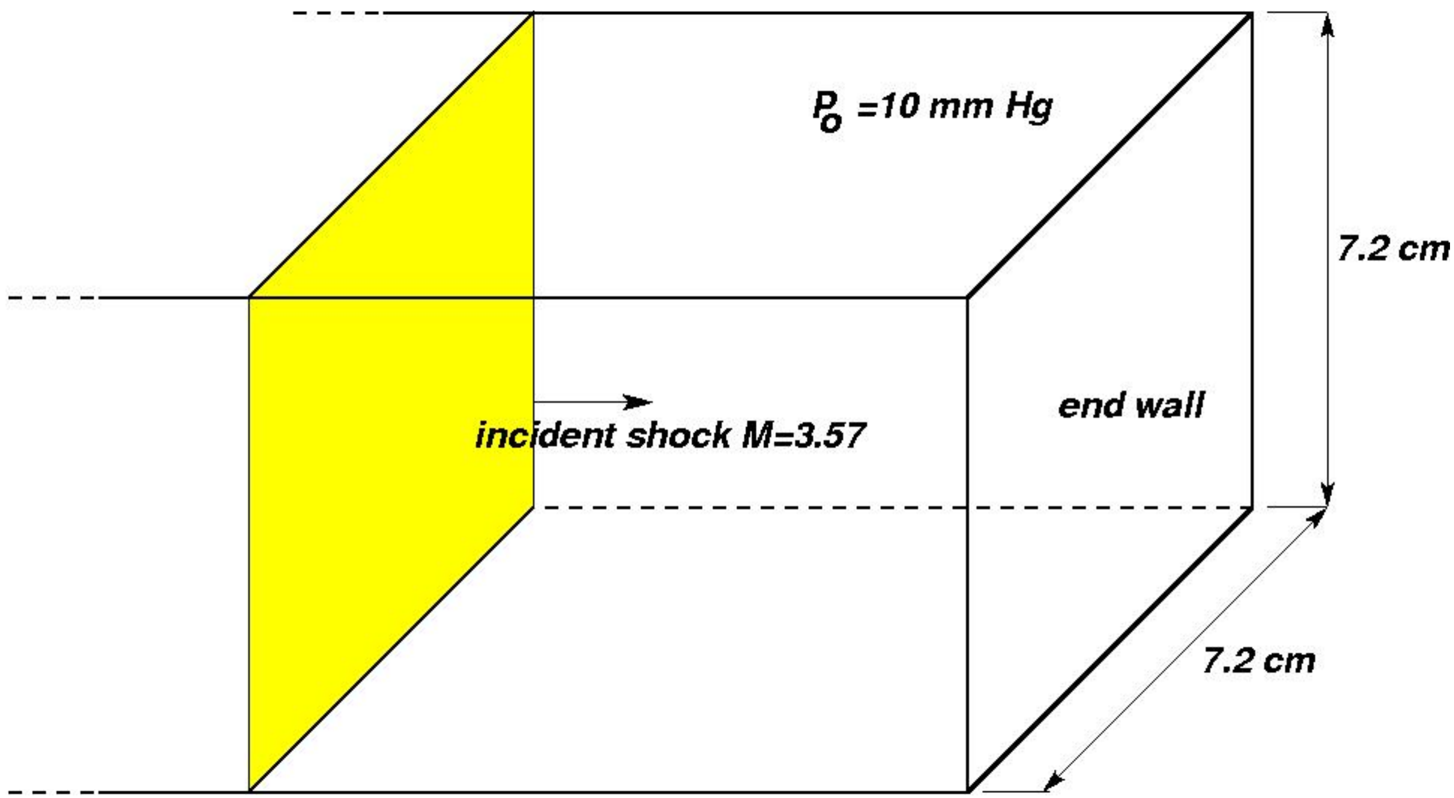
II. Simulations of reflected shock tube ignition time delay experiments in hydrogen-oxygen (strong and weak ignition) (~10M x 2)

III. Flame acceleration and DDT in hydrogen-oxygen in a long pipe. Predict run distance to detonation (~150M)

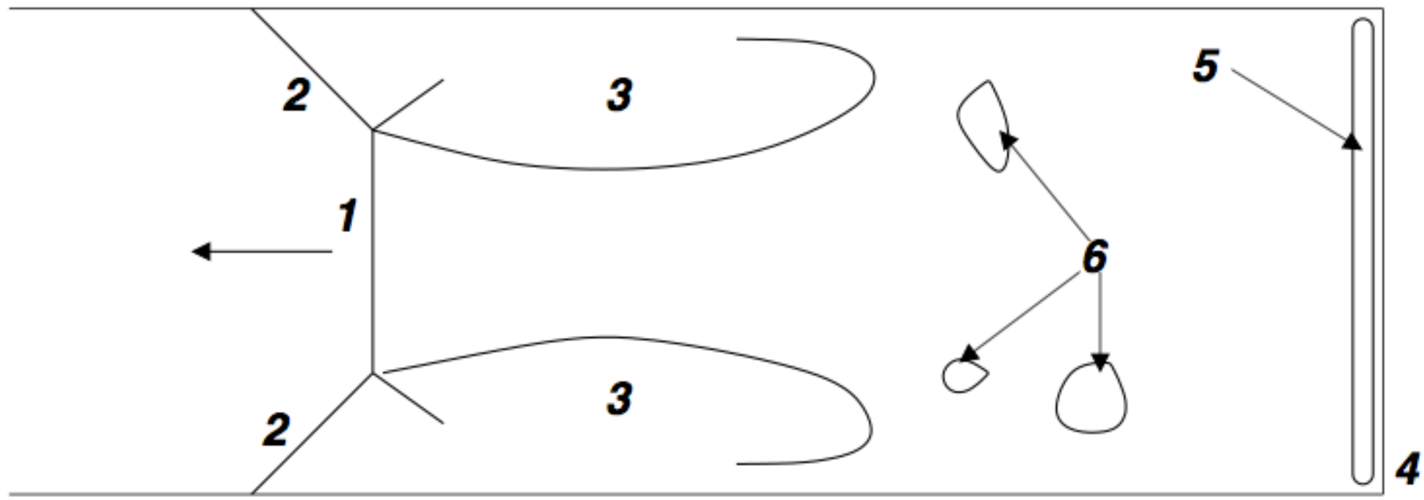
Setup and code features …



Setup figure for CO_2 configuration



Schematic shock bifurcation structure



1 Main reflected shock.

2 Oblique reflected shock (bifurcated foot).

3 Turbulent jet and recirculation bubble.

4 End wall.

5 A general location of strong ignition.

6 Weak ignition spots.

HSCD Code structure

Physics modules, initial/boundary conditions of a problem run on top of ALLA90.

ALLA90: a compressible Navier-Stokes fluid dynamics solver that runs on top of FTT.

FTT: fully threaded tree library provides mesh, adaptive mesh refinement (AMR), global parallel iterators, visualization, I/O.

HYDROGEN: physics and initial conditions and analysis/visualization.

HSCD (High-Speed Combustion and Detonation)

```
|
|--- ALLA90
|
|--- FTT
|
|--- HYDROGEN
```

HSCD code features

• 3D compressible reactive flow Navier-Stokes equations.

• H_2+O_2 kinetics: 8 species, 19 reactions.

• Multi-species NASA7 equation of state, temperature-dependent viscosity, heat conduction, mass diffusion, and radiation cooling.

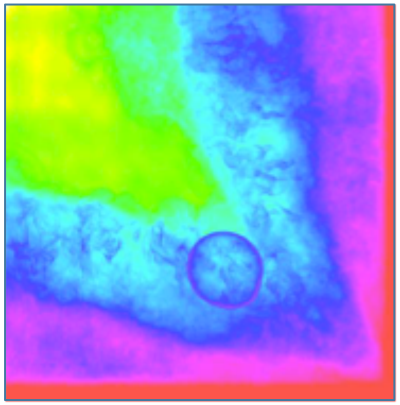
• Regular Cartesian mesh with cell-based AMR.

• Dynamic mesh refinement and mesh-re-balance every fourth time step.

• Written in Fortran 90 & C.

• MPI/OpenMP (added by Alexei Khokhlov and Charles Bacon but new algorithms had to be reinvented because decision mesh refinement stop scaling).

Scaling Challenges …



Scaling challenges I, BG/P

I/O

Moving from Lustre to GPFS required a different I/O strategy:

- Changed from one file per rank to a single file with MPI-IO.
- Improved checkpoint time by ~41x due to reduction in metadata overhead (28x faster) and due to MPI-IO enforcing aligned writes (~1.5x faster).
- Data acquired using the Darshan library.

Scaling challenges II, BG/P

Communications

- Original application was using a straight MPI.
- The code could not execute using 4 or 2 MPI ranks per node which lead to three idle cores in SMP mode.
- An AMR code (FTT) executes work-functions from the physical code (ALLA90) using global iterator provided by the FTT library which uses MPI.
- Putting OpenMP around work-functions resulted in a ~3x speedup of physical algorithms.
- This allowed scaling up to 32 racks on Intrepid.

Scaling challenges III, BG/P

Communications

- Reducing communication overheads.
- At this point the computational side of the code was scaling well.
- Efficiency loss at high rank counts was due to AMR decision-making algorithms and mesh re-balance.
- Performance was improved by ~100x on 128K cores by using new algorithms with one-sided communications.
- This reduced cost of AMR to ~10% of the cost of a run.

Current performance on BG/Q

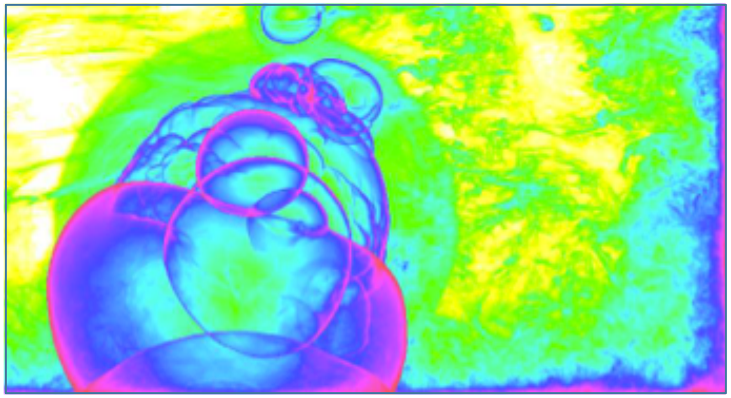
Moving the code to Mira resulted in a speedup of ~2.6x per core, ~9x per node.

Best performance is on 8 MPI ranks per node and 8 OMP threads per rank.

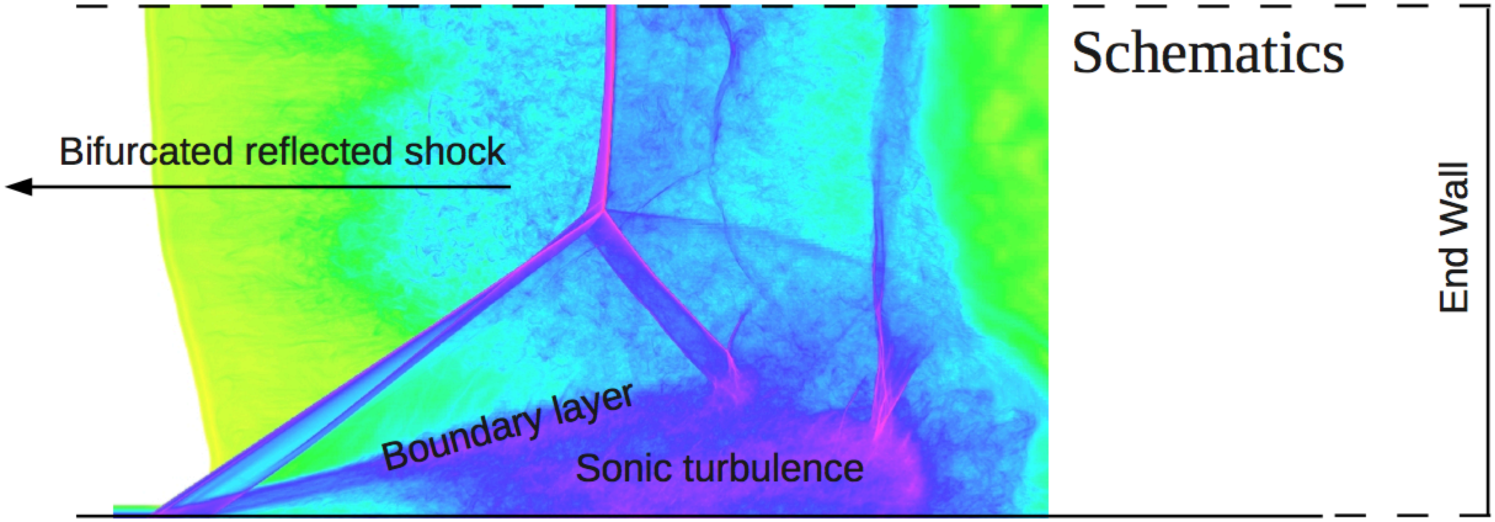
Compared to BG/P, cost of I/O went from <1% to ~20% of the total computational cost. The cause is unknown but it is currently under investigation and other features of the code are under study:

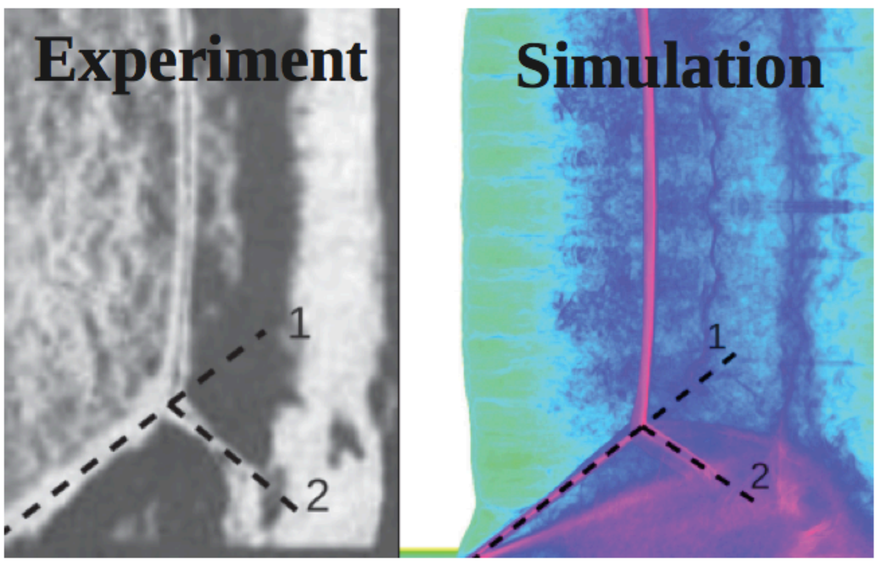
- Communications
- Data storage
- I/O
- Visualization
- node optimization
- Memory
- Data transfer
- ...

Simulations …



Reflected shock tube in CO_2 validation (BG/P)





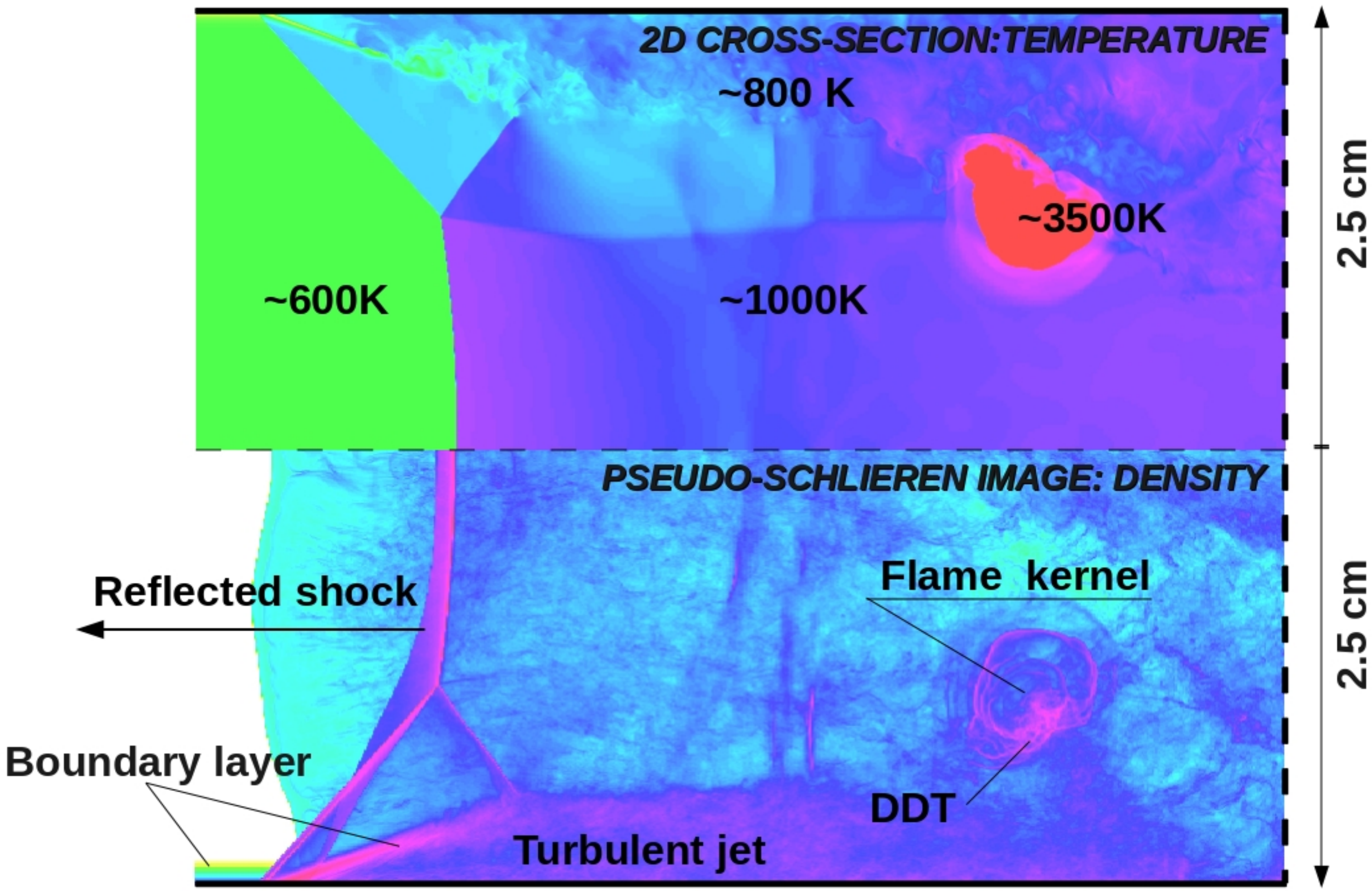
Comparison:

Angle	Experiment [*]	Simulation
1	36 deg	37 deg
2	-128 deg	-124 deg

[*] Brossard, J., Charpentier, N., Bazhenova, T.V. Fokeev, V.P., Kalachev, A., Kharitonov, A.I. Kharitonov, 15th Int. Symp. Shock waves & Shock Tubes, 163, 1984.

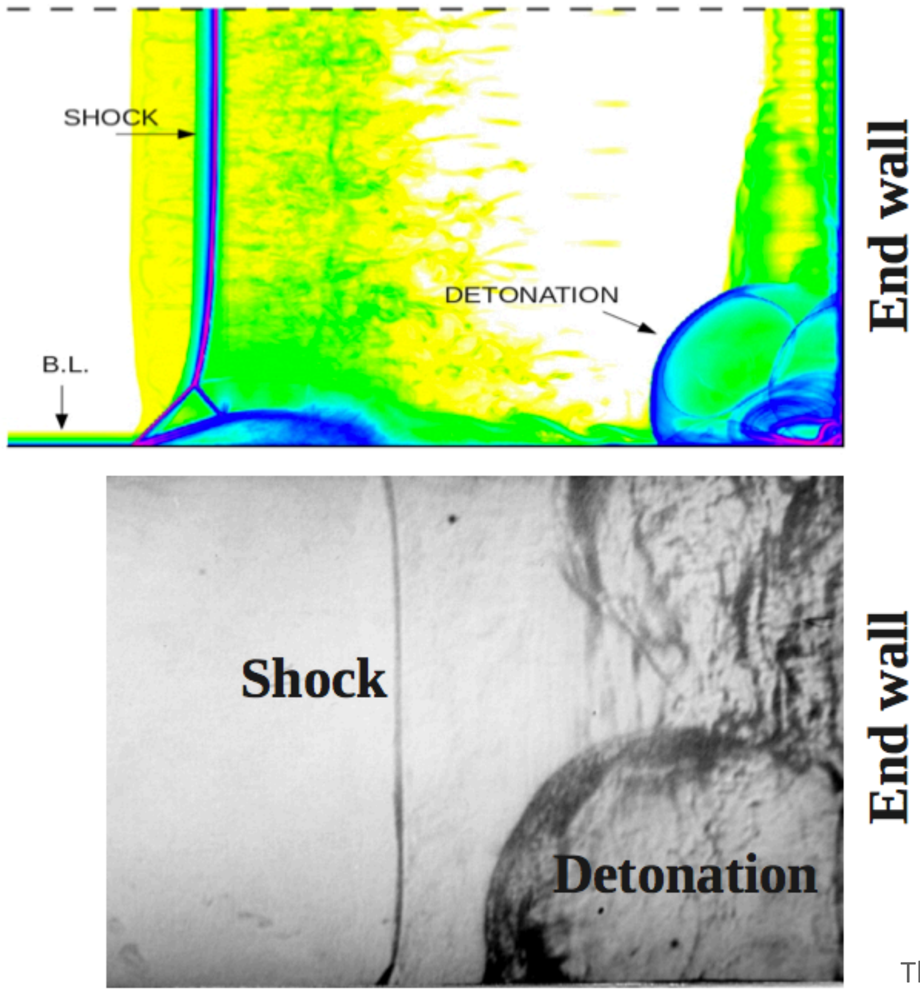
Simulation of <u>weak</u> ignition in $2H_2+O_2$ (BG/P)

incident Mach number = 1.47



Simulation of strong ignition in $2H_2+O_2$ (BG/P)

incident Mach number = 1.57



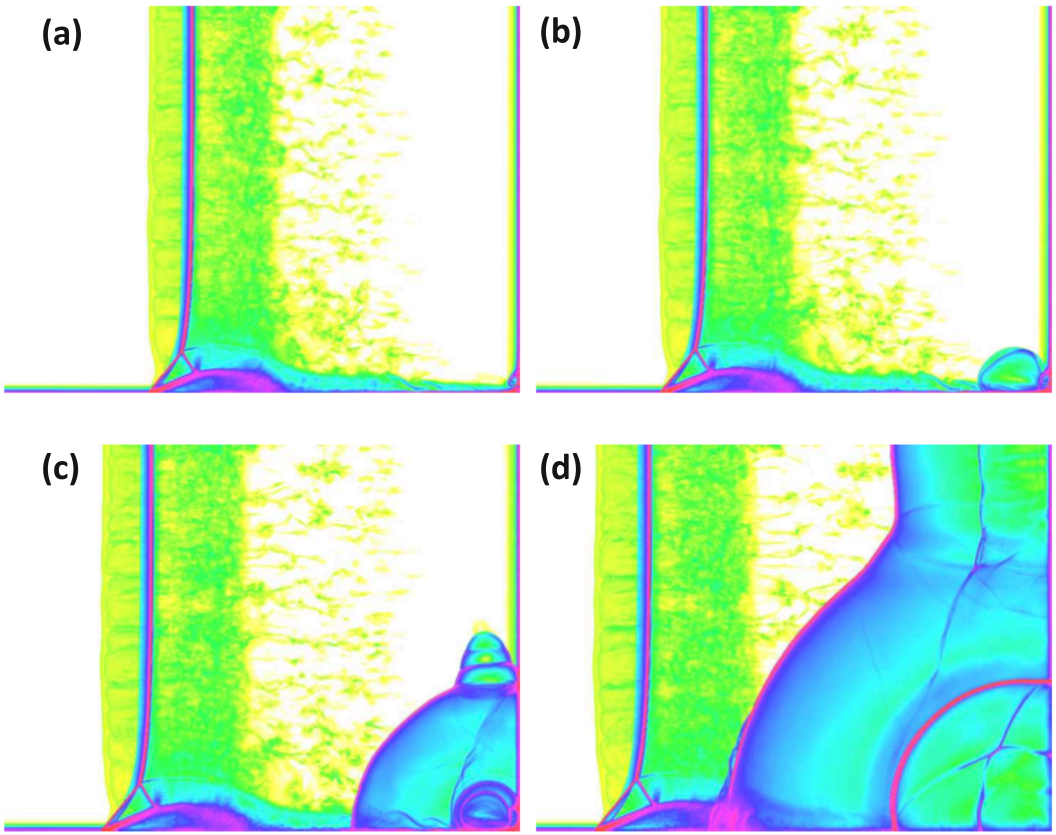
Thomas & Bambury 2001

Schlieren images of strong ignition in $2H_2+O_2$ (BG/P)

incident Mach number = 2.58

P = 0.1 atm
T = 300K

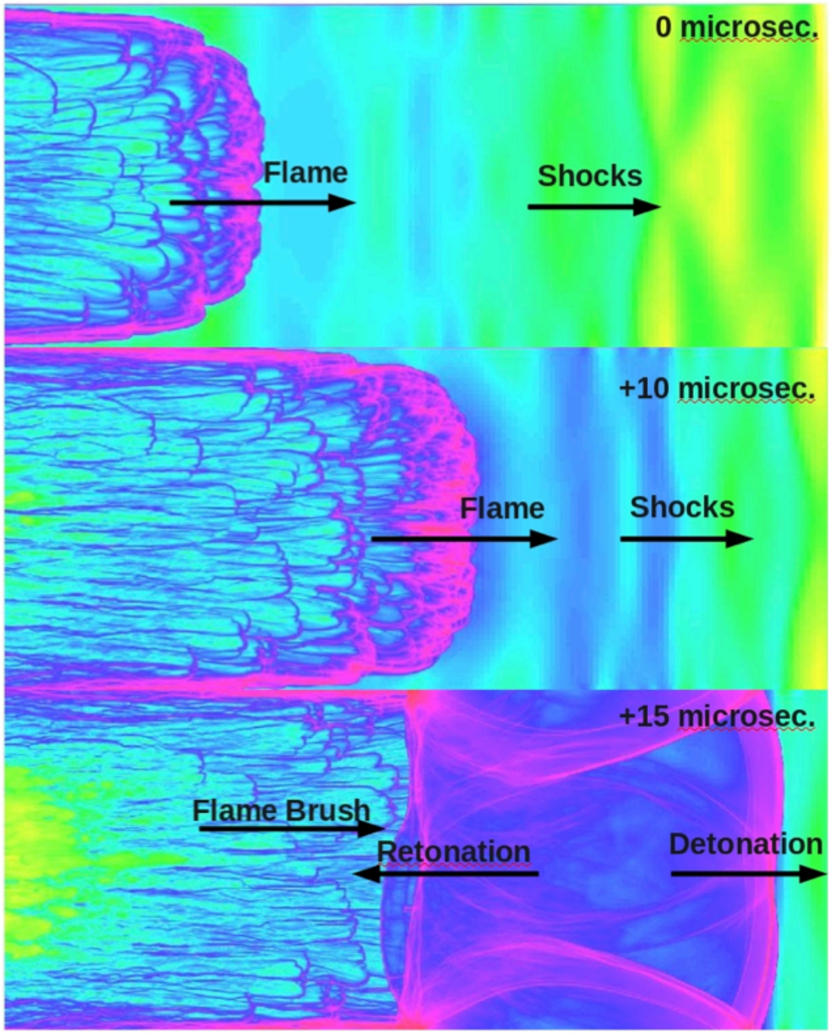
Ignition time
τ_i = 37 µsec.



(a) flow field immediately before an auto-ignition near the lower-right corner of the tube. The reflected bifurcated shock is on the left and moves to the left into the boundary layer which was created by the incident shock. The end wall is on the right. **(b)**, **(c)**, and **(d)** – development of a detonation in close proximity of the end-wall. Time difference between the frames is ≅ 0.3 µsec.

DDT in a hydrogen-oxygen mixture (BG/Q)

Preliminary, low resolution

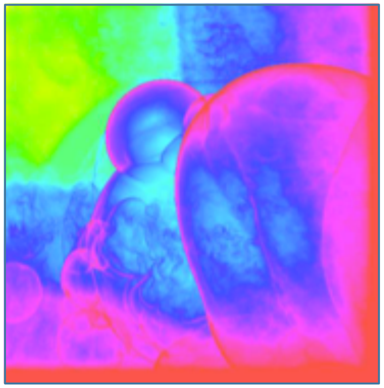


Next step:

High resolution simulations of DDT in a long pipe

Tube length: 172 cm
Cross-section: 2.7 cm x 2.7 cm
N cells ~ 10,000,000,000
N time steps ~ 50,000
Numerical resolution ~ 6 microns

Current timings …



Single-node scaling on BG/Q (VEAS)

Thread count	Timestep	Efficiency
1	349	100
2	190	92
4	110	79
8	70 (68)	62 (64)
16	64 (48)	34 (45)
32	45 (40)	24 (27)
64	50 (41)	11 (13)

Parenthetical numbers come after increasing the size of the array of cells passed to the work functions – high rank counts were getting not enough work per thread from the original setting

20 timesteps (includes 5x refine/balance)

BG/P
(Intrepid)

Node count	Time	Efficiency
512	2547	100
1024	1272	100
2048	716	88.9
4096	385	82.6

BG/Q
(VEAS)

Node count	Time	Efficiency
128	1430	100
256	716	99.9
512	390	91.6
1024	219	81.6

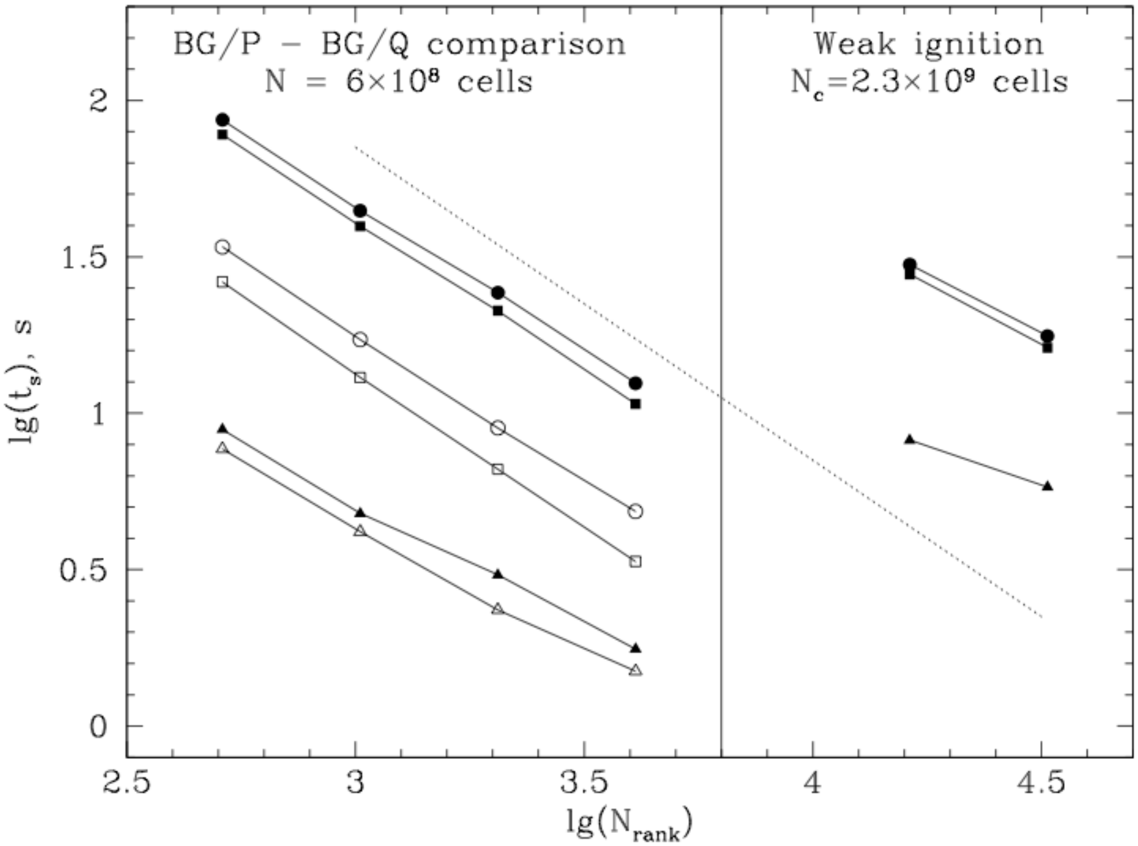
Main loop times

BG/P -> BG/Q speedup = 2.5x/core, 9.2x/node

BG/P Node count	Time	Efficiency
512	1779	100
1024	887	100
2048	488	91.1
4096	250	89

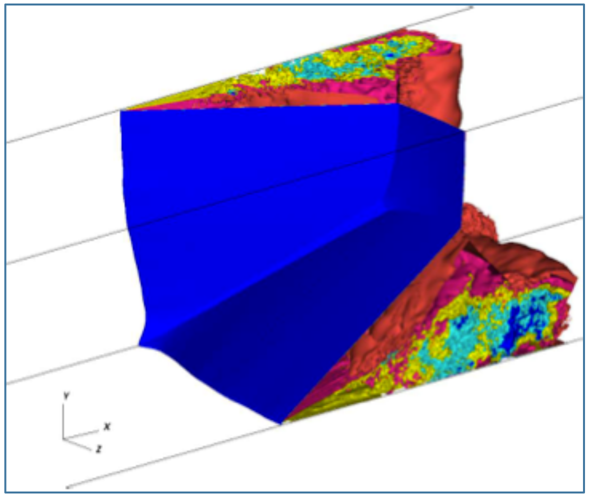
BG/Q Node count	Time	Efficiency
128	705	100
256	358	98.5
512	188	93.8
1024	101	87.3

Strong scalability of HSCD on BG/P & BG/Q



Strong scalability of HSCD code on BG/P Intrepid in MPI/OpenMP mode with 4 threads per rank (**black** symbols) and on BG/Q Vesta in MPI/OpenMP mode with 16 threads per rank and 4 ranks per node (open symbols). Squares - physics, triangles - AMR and load balance, circles -total. Time t_s is a wall-clock time of one time step. N_{rank} - number of MPI ranks. **Left** BG/P -BG/Q comparison. **Right** - data for a weak ignition case for the last hundred time steps of the simulation. Dashed line indicates ideal scaling.

Remaining challenges …



Remaining challenges

1. Solve I/O problems on BG/Q
2. Improve OpenMP performance of physical algorithms.
3. Due to increased number of OpenMP threads per rank used in the code we are beginning to see an overhead of integer calculations performed inside the FTT library on more than 64K ranks.
4. Parallelize FTT on each MPI rank using OpenMP.
5. Need new strategy for fine-grained OpenMP.
6. Help on visualization.

Questions …

?